Dr. **Albert Greenberg** (Uber), Title: "AI/ML at Uber: From Predictive to Generative Models".

There were some technical difficulties recording this course for UNITE so we lost the first part of the presentation (email said about 15 minutes but it may have been a bit longer according to some of the students I talked with). Therefore I went and read the abstract for the talk so I could have some idea as to what was going on. As a first semester student in the program it can be difficult enough to follow some of the discussions and even more so when you come into a presentation fairly far after the basic introduction.

The abstract states that Uber adopted AI way back in 2015. The use of AI has since evolved and any contact that a user has with Uber’s applications has AI/ML models working and interpreting the user’s input.

The lecture opens with Dr. Greenberg talking about some tool that Uber uses called Michelangelo. Since I missed the first part of the discussion I am wondering what this is exactly as it looks like a lot of different types of tools feed into it like Code, JIRA, Slack, Config, Code, etc.

We transition into a discussion on developer productivity by Wei(?) an engineer manager at Uber. It looks like they are simulating trips through AI which the developers can then use for testing. In addition, Uber Code Assistent uses AI to help developers code by making suggestions to them. Dr. Greenberg states that the senior developers want to stick with generative AI to help them test and suggest code modifications.

During the interactive metrics support they mentioned that they developed their own language for querying. Basically, their own dialect of SQL. (I wonder if they just translate into SQL behind the scenes). In general I do not like it when companies do this. It makes it more difficult to hire outside developers due to the proprietary nature of the technologies that are being used. The metrics assistant they created they call Genie and it mainly helps in consultation type questions, even customers use it. Dr. Greenberg states that Genie comes back with an answer 3/4s of the time. i.e. It has removed 3/4s of the work from the on-call team.

There is a point in the talk where Dr. Greenberg tells a story about AI telling the developers about a race condition in their code set. This is really cool and I wish I would have had such a tool when I was a code jockey. However, having an outside tool not only tell me about the race condition but also suggest a fix, I am at a loss for words.

Now a slide is being shown about generalist and task-specific LLMs (Large Language Models). I wonder if this presentation is about LLMs in general? I want the first 15 to 30 minutes of this presentation! However, the general AI strategy at Uber is to blend commercial AI models with open source models and put this on top of Uber’s proprietary platform (Michelangelo, Genie, etc). At the lowest level certain information needs to be checked to ensure compliance, privacy, and accuracy.

What Uber has learned is that they want to continue to use AI and to bet heavily on AI. They will use AI to improve developer productivity, end user experiences, and elsewhere in the company if they feel AI will provide a good return on investment. Uber feels this is a good strategy since AI platforms are evolving to serve a broader customer base. Lastly, LLMs are not quite where Uber wants them to be yet for an enterprise strategy.